

# PERANCANGAN DAN ANALISIS SISTEM SPEECH PROCESSING UNTUK TUNARUNGU MENGGUNAKAN METODE *HIDDEN MARKOV MODEL* DAN *MEL-FREQUENCY CEPSTRAL COEFFICIENT*

## *DESIGN AND ANALYSIS OF SPEECH PROCESSING SYSTEM FOR A DEAF PERSON USING HIDDEN MARKOV MODEL METHOD AND MEL-FREQUENCY CEPSTRAL COEFFICIENT*

Bagus Robbiyanto<sup>1</sup>, RaditianaPatmasari<sup>2</sup>, Rita Magdalena<sup>3</sup>

<sup>123</sup>Prodi S1 Teknik Telekomunikasi, Fakultas Teknik Elektro, Universitas Telkom, Bandung

<sup>1</sup>[bagusrobbi666@gmail.com](mailto:bagusrobbi666@gmail.com), <sup>2</sup>[raditiana@telkomuniversity.ac.id](mailto:raditiana@telkomuniversity.ac.id), <sup>3</sup>[magdalena@telkomuniversity.ac.id](mailto:magdalenat@telkomuniversity.ac.id)

### Abstrak

Mendengar merupakan salah satu cara untuk saling berkomunikasi, mendengar sangat dibutuhkan oleh manusia untuk mengerti maksud satu sama lain. Namun hal ini membatasi untuk orang normal berkomunikasi dengan tunarungu, karena tidak semua orang mengerti gerakan Bahasa isyarat.

Pada Tugas Akhir ini membuat dibuat suatu alat untuk membantu orang normal untuk berkomunikasi dengan orang yang menderita tunarungu. Alat ini mengolah sinyal suara input menjadi suatu *text* menggunakan metode *Mel Frequency Cepstral Coefficient* untuk mengekstrasi sinyal suara *input* dan diklasifikasi menggunakan metode *Hidden Markov Model* untuk melihat kemiripan antara sinyal suara yang sudah diekstrasi ciri dengan yang di *database*. Jika terdapat suatu kemiripan maka menghasilkan suatu *text*, kemudian *text* tersebut diolah menjadi suatu *input* baru yang menampilkan video Bahasa Isyarat Indonesia.

Hasil pengujian menunjukkan bahwa kombinasi metode *mel frequency cepstral coefficient* dan *Hidden Markov Model* mampu mengenali sinyal suara berupa kata dengan akurasi tertinggi mencapai 87%.

**Kata Kunci:** Bahasa Isyarat indonesia, Tunarungu, MFCC, HMM.

### Abstract

*Hearing is one way to communicate with each other, hearing is needed by manusia to understand each other's intentions. But this limits the normal person communicating with the deaf, because not everyone understands Sign Language.*

*In this Final Project, a tool is created to help normal people communicate with people who are deaf. This tool processes the input sound signal into a text using the Mel Frequency Cepstral Coefficient method to extract input sound signals and is classified using the Hidden Markov Model method to see the similarity between the sound signals that have been extracted and those in the database. If there is a similarity then it produces a text, then the text is processed into a new input that displays Indonesian Sign Language videos.*

*The test results showed that the combination of the mel frequency cepstral coefficient method and the Hidden Markov Model were able to recognize sound signals in the form of words with the highest accuracy reaching 87%.*

**Keywords:** Indonesian Sign Language, Deaf Person, speech processing, MFCC, HMM.

## 1. Pendahuluan

### 1.1 Latar Belakang Masalah

Tunarungu (*a deaf sperson*) adalah orang yang mengalami ketidakmampuan mendengar, sehingga mengalami hambatan dalam memproses informasi bahasa melalui pendengarannya dengan atau tanpa menggunakan alat bantu dengar (*hearing aid*) [1].

Komunikasi dibagi menjadi 2 jenis yaitu komunikasi verbal dan non verbal. Komunikasi verbal yaitu komunikasi yang disampaikan secara tertulis atau berbicara dan komunikasi non verbal yaitu komunikasi yang tidak menggunakan kata-kata [2].

Bahasa isyarat unik dalam jenisnya di setiap negara. Bahasa isyarat bisa saja berbeda di negara-negara yang berbahasa sama. Untuk Indonesia, sistem yang sekarang umum digunakan adalah Sistem Isyarat Bahasa Indonesia (BISINDO) yang sama dengan bahasa isyarat *America (ASL -American Sign Language)*. Sistem Isyarat Bahasa Indonesia (BISINDO) yang baku merupakan salah satu media bantu tuna wicara dalam bermasyarakat. Wujudnya adalah tatanan sistematis tentang seperangkat isyarat jari, tangan, dan berbagai gerak yang melambangkan kosakata Bahasa Indonesia [3].

Pada penelitian sebelumnya [4] sudah dilakukan penelitian yang berjudul "*Design Implementasi Voice Command*" menggunakan metode MFCC dan HMM dengan akurasi yang didapatkan 93,89% pada lingkungan

tanpa noise dan 58,1% pada lingkungan noise. Namun pada penelitian sebelumnya hanya dilakukan pengolahan suara yang di proses lalu menghasilkan kata.

Pada penelitian ini akan dibuat sebuah sistem yang akan membantu orang normal agar bisa berkomunikasi dengan tunarungu melalui sebuah video gerakan bahasa isyarat. Input dari sistem ini berupa sinyal suara yang di proses menggunakan metode *Hidden Markov Model* dan ekstrasi ciri *Mel-Frequency Cepstral Coefficient* Sehingga menghasilkan kata selanjutnya kata tersebut dijadikan input untuk memanggil output video gerakan bahasa isyarat.

## 2. Dasar Teori

### 2.1 Bahasa Isyarat Indonesia (BISINDO)

Bahasa isyarat pada dasarnya sama dengan bahasa lisan akan tetapi cara penyampaiannya yang tidak sama, bahasa isyarat menggunakan tangan dan bahasa lisan menggunakan mulut. Sebelumnya bahasa isyarat tidak mendapat tempat dimasyarakat pengguna Bahasa Indonesia, sehingga bahasa isyarat menjadi tidak jelas dan tidak menentu.

Bahasa Isyarat Indonesia adalah isyarat-isyarat kata yang pada mulanya diambil dari isyarat-isyarat yang disampaikan anak tunarungu yang bisa diterima sebagai kata atau kosakata dalam Bahasa Indonesia termasuk *America Sign Language* (ASL) atau Bahasa Isyarat Amerika yang diubah menjadi Bahasa Indonesia[5].

### 2.2 Sinyal Ucapan Manusia

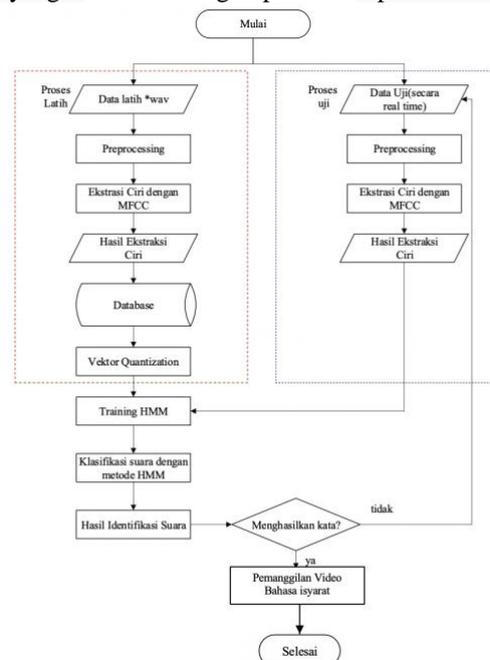
Proses pembentukan ucapan manusia disebabkan oleh suatu produksi yang dilakukan oleh organ tubuh manusia. Proses tersebut membutuhkan kerja sama organ tubuh manusia sehingga dapat mengeluarkan sinyal ucapan. Dimulai dengan memberikan pesan kepada otak pembicara, lalu pesan tersebut diubah menjadi perintah yang dikirimkan kepada masing masing organ tubuh manusia. Jika diamati sinyal ucapan manusia dapat berubah terhadap waktu. Tetapi karakteristik sinyal bersifat tetap pada selang waktu (5 sampai 100 mili detik) untuk mengetahui karakteristik sinyal berubah-ubah dibutuhkan selang waktu lebih panjang [7].

### 2.3 Automatic Speech Recognition (ASR)

*Speech Recognition System* atau disingkat ASR adalah suatu sistem yang berfungsi menerjemah suatu bahasa lisan menjadi bentuk data komputer. Sistem ini menggunakan mikrofon untuk menerima informasi lalu membandingkan dengan database yang tersedia.

Secara konseptual sistem ini bekerja dengan mengubah kata-kata menjadi angka-angka yang berbentuk tulisan lalu membandingkan dengan kode yang terdapat pada database untuk melakukan perintah suatu pekerjaan [8].

Pada tugas akhir ini akan dirancang sebuah sistem untuk mengidentifikasi sinyal suara berupa huruf *alphabet* menggunakan ekstrasi ciri *mel frequency cepstral coefficient* dan klasifikasi menggunakan *Hidden Markov Model*. Adapun diagram alir yang akan dirancang dapat dilihat pada Gambar 2.6.



Gambar 2.6 Diagram alir sistem

## 2.4 Mel Frequency Cepstral Coefficient

MFCC merupakan cara yang paling sering digunakan pada berbagai bidang area pemrosesan suara, karena dianggap cukup baik dalam merepresentasikan sinyal. MFCC menganut cara kerja telinga manusia dimana telinga manusia merupakan filter linier pada frekuensi rendah dan bekerja secara logaritmik pada frekuensi tinggi. Metode MFCC menggunakan 2 filter yaitu filter linier dibawah 1000 Hz dan logaritmik diatas 1000Hz nilai-nilai frekuensi tersebut dengan frekuensi *mel* [9][10].

### 2.4.1 Frame Blocking

Sinyal suara manusia merupakan sinyal yang tidak stabil. Oleh karena itu dibutuhkan *frame blocking* untuk memotong sinyal menjadi bagian yang lebih kecil sehingga didapat karakteristik suara yang stabil. Biasanya sinyal suara di potong setiap 25ms.

### 2.4.2 Windowing

Proses *windowing* dilakukan untuk mengurangi kebocoran spektra yang merupakan efek dari timbulnya sinyal baru yang memiliki frekuensi yang berbeda dari frekuensi aslinya. Hal tersebut dapat terjadi karena rendahnya *sampling rate* atau karena proses *framing* yang menyebabkan sinyal menjadi *diskontinu*. Bila *window* didefinisikan.

$$y(n) = x(n)w(n), 0 \leq n \leq N - 1 \quad (2.1)$$

### 2.4.3 Fast Fourier Transform

Untuk mendapat sinyal dalam domain frekuensi salah satu metode yang dapat digunakan adalah DFT (*Discrete Fourier Transform*). DFT dilakukan terhadap semua *frame* yang telah di-*windowing* sebelumnya. Untuk mempercepat proses DFT dan mengurangi beban komputasi maka dilakukanlah FFT (*Fast Fourier Transform*) yaitu metode transformasi *Fourier* dengan proses lebih cepat.

$$f(n) = \sum_{k=0}^{N-1} y_k e^{-2\pi jkn/N}, n = 0, 1, 2, \dots, N - 1 \quad (2.2)$$

### 2.4.4 Mel Frequency Wrapping

Pada proses ini dilakukan pengubahan skala frekuensi menjadi skala Mel. Skala *Mel-Frequency* adalah frekuensi linier di bawah 1KHz dan Logaritmik diatas 1KHz. Skala Mel dapat diperoleh dengan pendekatan persamaan :

$$mel(f) = 1125 * \log_{10}(1 + \frac{f}{700}) \quad (2.3)$$

### 2.4.5 Cepstrum

Selanjutnya dilakukan DCT (*Discrete Cosine Transform*) untuk memperoleh sinyal dalam domain frekuensi. Hasil keluaran dari proses DCT ini disebut *Mel-Frequency Cepstral Coefficient (MFCC)*. Pada sistem pengenalan suara biasanya hanya diambil 12 atau 13 koefisien pertama dari MFCC untuk mendapat hasil ekstraksi yang baik. MFCC bias didapatkan dari persamaan:

$$cn = \sum_{k=1}^K (\log Sk) \cos[n(k - \frac{1}{2}) \frac{\pi}{K}], n = 1, 2, \dots, K \quad (2.4)$$

## 2.5 Hidden Markov Model

*Hidden Markov Model* berdasarkan deret observasi dapat dibedakan atas *Diskrit HMM (DHMM)* dan *Continous Density HMM (CDHMM)*. Pada penelitian ini digunakan *Continous Density Hidden Markov Model (CDHMM)* adalah sebuah model statistik dari sebuah sistem yang diasumsikan sebuah proses dengan parameter yang tak diketahui, dan tantangannya adalah menentukan parameter-parameter tersembunyi (*hidden state*) dari parameter-parameter yang dapat diamati (*observed state*). Metode HMM akan digunakan untuk proses pengklasifikasian ciri sistem *speech recognition*. Metode ini digunakan untuk pelatihan dan pengujian sistem mengenali ciri sinyal suara dari parameter observasi yang ada [11].

Sebuah Hidden Markov Model dikarakteristikan dengan parameter berikut (Rabiner, 1989)[12]:

HMM diskrit terdiri dari lima komponen, yaitu:

- Jumlah *state* (N)
- Parameter model (M)
- Intial state* atau *state* awal  $\pi = \{\pi_i\}$
- Probabilitas transisi (A)
- Probabilitas simbol observasi

### 2.5.1 Vector Quantization

*Vector Quantization* diperlukan untuk menghasilkan deretan observasi (indeks *codebook*) yang selanjutnya akan diproses menggunakan *Hidden Markov Model (HMM)* untuk proses training HMM. Sebelum suatu vektor *feature* dikuantisasi, terlebih dahulu harus dibentuk vektor *codebook* yang akan mengkuantisasi vektor *feature* tersebut. *Codebook* dibentuk menggunakan algoritma *clustering*.

Pada kuantisasi vektor ini, algoritma *clustering* yang digunakan adalah K-Means *clustering*. Dasar algoritma K-means adalah sebagai berikut [11]:

- Tentukan nilai k sebagai jumlah cluster yang ingin dibentuk.

- b. Bangkitkan  $k$  *centroid* (titik pusat cluster) awal secara random.
- c. Hitung jarak setiap data ke masing-masing *centroid* menggunakan rumus korelasi antar dua objek, yaitu *Euclidean Distance* dan kesamaan *Cosine*.
- d. Kelompokkan setiap data berdasarkan jarak terdekat antara data dengan *centroidnya*.
- e. Tentukan posisi *centroid* baru ( $k$  C) dengan cara menghitung nilai rata-rata dari data-data yang ada pada *centroid* yang sama.

$$c_k = \left( \frac{1}{n_k} \right) \sum d_i \quad (2.8)$$

Dimana  $n_k$  adalah jumlah dokumen dalam cluster  $k$  dan  $d_i$  adalah dokumen dalam cluster  $k$ .

- f. Kembali ke langkah 3 jika posisi centroid baru dengan centroid lama tidak sama.

### 2.5.2 Training HMM

Data yang digunakan pada proses *training* adalah data berupa indeks hasil *vector quantization*, indeks ini dapat dikatakan sebagai simbol observasi *Hidden Markov Model*. Proses *training* menggunakan metode *Hidden Markov Model (HMM)*. Pemodelan HMM terdiri dari 3 matriks probabilitas, yaitu matriks transisi antar *state* (A), matriks probabilitas pengamatan suatu *state* (B), dan matriks probabilitas awal *state* ( $\pi$ ), jenis HMM yang digunakan merupakan diskrit ergodic, dimana parameter-parameter HMM seperti matriks A, B, dan  $\pi$  dibangkitkan secara *random* dengan nilai yang dinormalisasi ke satu. Nilai-nilai matriks A, B, dan  $\pi$  tersebut kemudian dilakukan re-estimasi melalui proses pelatihan untuk mendapatkan nilai parameter yang optimal.

Sedangkan untuk elemen-elemen pembentuk HMM yang lain diambil berdasarkan banyaknya *state* (N) dan jumlah observasi (M). Pada HMM *state*-nya bersifat *hidden*, dimana *state* yang *hidden* tersebut adalah jenis sinyal suara itu sendiri, sedangkan bagian yang akan diobservasi adalah ciri dari sinyal suara. Re-estimasi parameter HMM merupakan hasil implementasi dari algoritma *Baum-Welch* atau *Expectation Maximum (EM)*. Hasil dari re-estimasi parameter HMM adalah nilai baru elemen-elemen matriks A, B, dan  $\pi$ . Iterasi untuk re-estimasi dihentikan apabila iterasi sudah maksimal atau jika model baru tidak memberikan perbaikan yang cukup berarti.

$$\log(p(O|\lambda)) - \log(p(O|\lambda)) \leq 0.001 \quad (2.9)$$

Dalam mengoperasikan algoritma HMM, perlu diperhatikan mengenai besarnya nilai-nilai parameter yang terjadi pada proses perhitungan. Hal ini penting karena nilai parameter yang sangat kecil dapat menyebabkan nilai tersebut dianggap nol serta jika digunakan sebagai faktor pembagi maka akan diperoleh harga yang sangat besar. Oleh karena itu digunakan penskalaan pada perhitungan parameter-parameter dalam model HMM sehingga diperoleh algoritma baru untuk menghitung  $\alpha_t(i)$  dan  $\beta_t(i)$ . Faktor skala ( $C_t$ ) yang digunakan adalah:

$$C_t = \frac{1}{\sum_{i=1}^N \alpha_t(i)} \quad (2.10)$$

### 2.5.3 Algoritma Baum Welch

Pada algoritma *Baum-Welch* sangat mendukung dalam tahap *training* menggunakan *multiple observation sequence* sehingga sangat sesuai dengan permasalahan yang dihadapi. Adapun langkah-langkah dari algoritma *Baum-Welch* yaitu [13]:

- a) Prosedur *forward*

pada prosedur *forward* dapat didefinisikan dengan:

$$\alpha_t(i) = P(O_1, O_2, \dots, O_t, i_t = i | \lambda) \quad (2.11)$$

$\alpha_t(i)$  dapat dihitung sebagai berikut :

$$\alpha_t(i) = \pi_i b_i(O_t) \quad (2.12)$$

$$\alpha_{t+1}(i) = b_j(O_{t+1}) \sum_{i=1}^N \alpha_t(i) \cdot \alpha_{ij}$$

- b) Prosedur *backward*

pada prosedur *backward* dapat didefinisikan dengan:

$$\beta_t(i) = P(O_{t+1}, O_{t+2}, \dots, O_T, i_t = i | \lambda) \quad (2.13)$$

$$\beta_T(i) = 1, 1 \leq i \leq N \quad (2.14)$$

$$\beta_t(i) = b_j(O_{t+1}) \sum_{i=1}^N \beta_{t+1}(i) \cdot \alpha_{ij}$$

## 2.7 Performansi Sistem

Sistem yang sudah dirancang akan diuji tingkat akurasinya dalam menghasilkan *output* yang sesuai dengan yang diinginkan. Untuk perhitungan akurasi sistem adalah sebagai berikut:

$$akurasi = \frac{jumlah\ data\ benar}{jumlah\ data\ yang\ diuji} \times 100\% \quad (2.7)$$

## 3. PENGUJIAN DAN ANALISIS

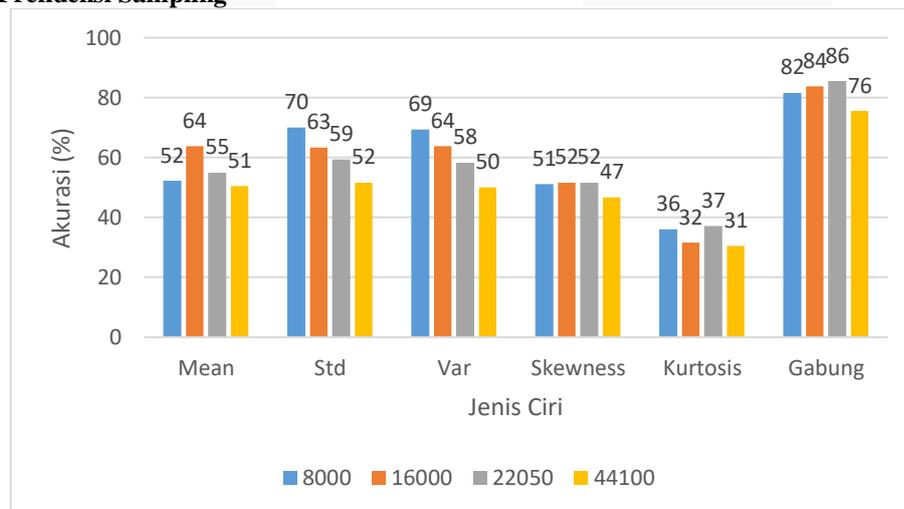
### 3.1 Pengujian Sistem

Pengujian sistem pada alat bantu belajar Bahasa isyarat ini dilakukan untuk mengukur performansi dan Pada bab ini dilakukan pengujian untuk menguji performansi dari sistem yang telah dirancang pada bab III. Sistem *speech processing* untuk tuna rungu menggunakan metode ekstraksi ciri MFCC dan klasifikasi HMM. Pengujian terdiri dari 5 skenario yang terdiri dari:

1. Pengaruh frekuensi sampling terhadap akurasi sistem *speech processing*.
2. Pengaruh ukuran frame terhadap akurasi sistem *speech processing*.
3. Pengaruh jumlah koefisien MFCC pada proses ekstraksi ciri terhadap akurasi sistem *speech processing*.
4. Pengaruh jumlah data latih tiap kelas atau data set pada pembangunan model klasifikasi HMM terhadap akurasi sistem *speech processing*.
5. Pengaruh iterasi pelatihan HMM terhadap akurasi sistem *speech processing*.

Pada tugas akhir ini, total data audio yang digunakan sebanyak 450 data rekaman audio yang terdiri dari 30 kelas kata. Tiap kelas kata terdiri dari tiga orang dengan masing-masing orang mengucapkan 5 kali.

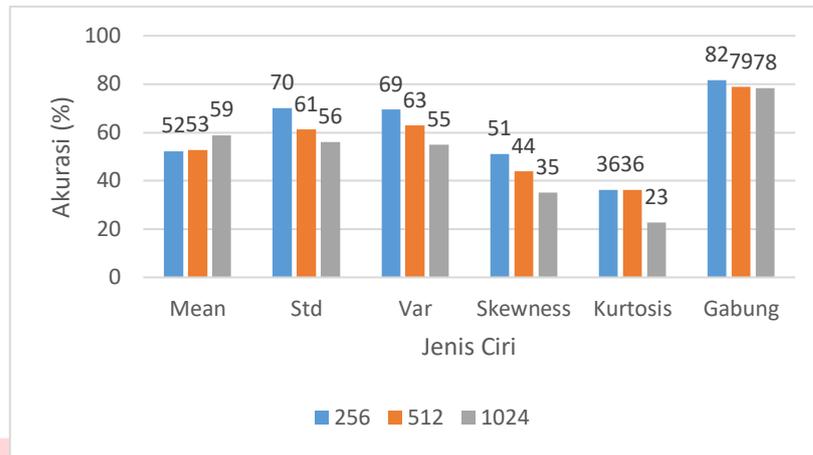
#### 3.2.1 Pengaruh Frekuensi Sampling



**Gambar 3.1** Grafik Pengaruh Frekuensi Sampling

menunjukkan nilai akurasi ciri gabung menghasilkan akurasi yang paling tinggi dibandingkan dengan jenis ciri lainnya. Akurasi tertinggi terdapat pada frekuensi sampling 22050 sampel/detik. Pada ciri gabung, semakin besar frekuensi sampling menyebabkan nilai akurasi semakin tinggi saat perubahan frekuensi sampling 8000, 16000, dan 22050. Pada frekuensi sampling 44100, nilai akurasi pada semua jenis ciri menjadi turun.

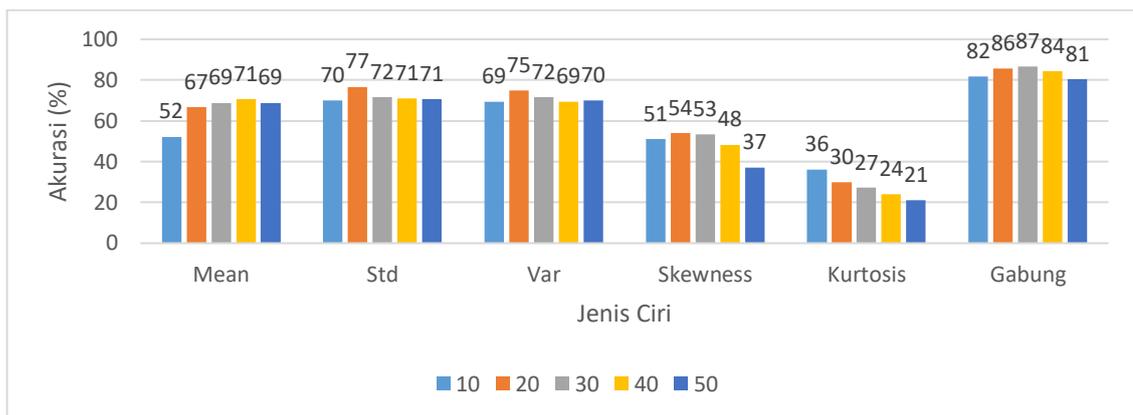
#### 3.2.2 Analisis pengaruh ukuran frame



Gambar 3.2 Grafik Pengaruh Ukuran Frame

Berdasarkan gambar 3.2 hasil pengujian menunjukkan nilai akurasi ciri gabung menghasilkan akurasi yang paling tinggi dibandingkan dengan jenis ciri lainnya. Akurasi tertinggi terdapat pada ukuran frame sebesar 256 sampel data. Pada ciri gabung, semakin besar ukuran frame menyebabkan nilai akurasi semakin rendah atau berbanding terbalik. Pada ciri mean, nilai akurasi meningkat saat nilai ukuran frame membesar atau berbanding lurus, namun akurasi pada ciri mean masih rendah yaitu di bawah 60%. Ukuran frame terbaik pada pengujian ini adalah 256 karena akurasi paling tinggi dan membuktikan bahwa nilai ciri MFCC lebih baik menggunakan ukuran frame yang kecil. Ukuran frame yang kecil menghasilkan analisis frekuensi yang lebih detail dan menghasilkan model klasifikasi HMM yang lebih baik dengan akurasi 82% pada ciri gabung.

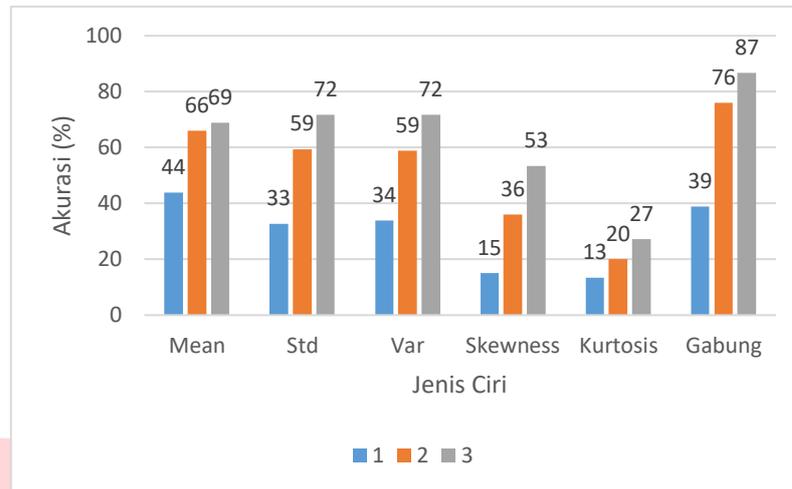
3.2.3 Analisis pengaruh koefisien MFCC



Gambar 3.3 Hasil Analisis Pengaruh Koefisien MFCC

Berdasarkan gambar 3.3 Akurasi terbaik didapatkan ketika melakukan pengambilan data di tempat A yaitu dengan akurasi sebesar 87,3%. Hal ini terjadi dikarenakan dalam pengambilan data uji menggunakan tempat dan kondisi yang sama dengan tempat pengambilan data latih akan memungkinkan dalam pengambilan data suara akan dimiliki model suara yang lebih mirip antara data latih dengan data uji, sedangkan kondisi tempat B dan C memiliki noise yang membuat sistem salah dalam mengenali suara.

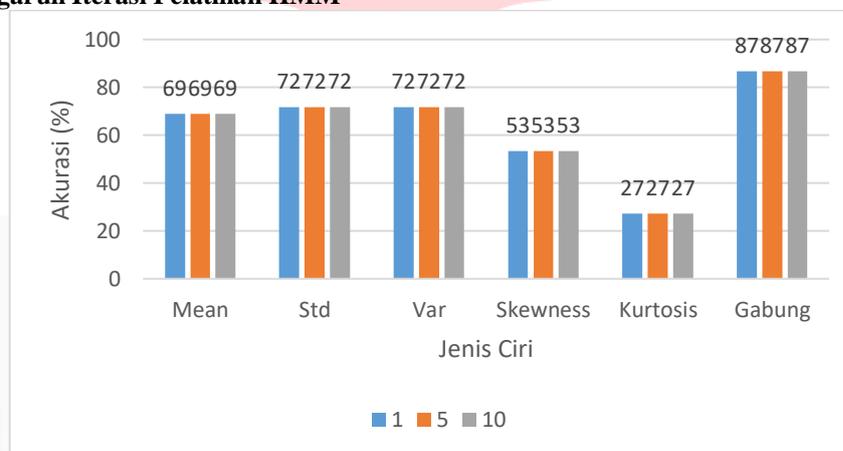
3.2.4 Analisis Pengaruh Data Set



**Gambar 3.4** Analisis Pengaruh Data Set

Dari gambar 3.4 menunjukkan semakin besar nilai data set maka semakin besar pula nilai akurasi untuk semua jenis ciri statistik. Hasil ini berarti semakin banyak data set maka model klasifikasi HMM lebih baik dalam proses pelatihannya untuk menghasilkan probabilitas tiap kelas kata yang lebih baik.

### 3.2.5 Analisis Pengaruh Iterasi Pelatihan HMM



**Gambar 3.4** Analisis Pengaruh Iterasi Pelatihan HMM

Dari gambar 3.5 Pengujian iterasi HMM menghasilkan hasil yang sama untuk iterasi yang semakin besar. Hasil ini berarti untuk jenis ciri MFCC pada kasus suara rekaman kata pada tugas akhir ini tidak memerlukan jumlah iterasi yang banyak karena hanya menggunakan satu kali iterasi hasil sudah baik. Nilai iterasi diperbesar tidak mengubah besarnya akurasi berarti semakin banyak perhitungan proses *baum welch* tidak mengubah parameter klasifikasi HMM menunjukkan nilai akurasi ciri gabung menghasilkan akurasi yang paling tinggi dibandingkan dengan jenis ciri lainnya. Akurasi tertinggi sebesar 87% terdapat pada pada ciri gabung.

## 4. Kesimpulan dan Saran

### 4.1 Kesimpulan

Model *speech recognition* yang mampu mengubah dan mendeteksi sinyal suara dalam bahasa Indonesia menjadikannya suatu teks telah dirancang. Hasil kesimpulan untuk tugas akhir ini adalah akurasi sistem terbaik sebesar 87%. Parameter terbaik yaitu frekuensi sampling sebesar 8000 sampel/detik, ukuran frame sebanyak 256 data per frame, jumlah koefisien MFCC sebanyak 30 koefisien, jumlah data set sebanyak 3 data per orang tiap kelas kata, dan iterasi pelatihan HMM sebanyak 1 kali. Frekuensi sampling sinyal audio cukup menggunakan 8000 sampel/detik sesuai dengan teori. Ukuran frame yang kecil menghasilkan ciri MFCC yang lebih detail dan menghasilkan model klasifikasi HMM yang terbaik untuk 30 kelas kata. Akurasi terbaik dicapai pada pengujian menggunakan data sebanyak 450 data sinyal rekaman. Ratio perbandingan data latih dan data uji adalah 9:6 sehingga jumlah data latih sebanyak 270 data latih dan 180 data uji.

## 4.2 Saran

Berdasarkan hasil pengujian dan analisis yang telah dilakukan, sistem yang terhubung ke device ini masih mampu untuk dikembangkan menjadi lebih baik dan akurat dalam mengidentifikasi suatu sinyal suara berupa huruf alphabet. Adapun saran untuk pengembangan sistem yang sudah dibuat ini yaitu: 1. Device yang digunakan masih berukuran besar. 2. Dilakukan penelitian menggunakan metode ekstraksi ciri dan klasifikasi lainnya sebagai pembandingan untuk melihat akurasi terbaik. 3. Dilakukan penelitian dengan mengembangkan data input, yaitu tidak hanya huruf alphabet, namun dapat data input dapat dikembangkan menjadi kata bahkan kalimat. 4. Dilakukan pengambilan data latih yang lebih banyak. Data latih diambil dengan berbagai tempat dan berbagai jenis suara. 5. Dibutuhkan suatu filter untuk meningkatkan akurasi diruangan yang tidak kedap suara.

## DAFTAR PUSTAKA

- [1] T. Hernawati, "Pengembangan kemampuan berbahasa dan berbicara anak tunarungu," vol. 7, 2007.
- [2] A. M. Hardjana, *Komunikasi Interpersonal dan Intrapersonal*. Yogyakarta: Kanisius, 2003.
- [3] Sutarman, M. A. Majid, and J. M. Zain, "A review on the development of Indonesian sign language recognition system," *J. Comput. Sci.*, vol. 9, no. 11, pp. 1496–1505, 2013.
- [4] M. Sidiq, T. A. B. W, and S. Sa'adah, "Desain dan Implementasi Voice Command Menggunakan Metode MFCC dan HMMs," *Epoch*, vol. 2, no. 1, pp. 1-10, 2012.
- [5] M. Hendra Pradikja, H. Tolle, and K. Candra Brata, "Pengembangan Aplikasi Pembelajaran Bahasa Isyarat Berbasis Android Tablet," vol. 2, no. 8, pp. 2548–964, 2018.
- [6] M. Najiburahman, "Simulasi dan Analisis Sistem Penerjemah Bahasa SIBI Menjadi Bahasa Indonesia Menggunakan Metode Klasifikasi Hidden Markov Model," *e-Proceeding of Engineering .*, vol. 2, no. 1. pp. 97–105, 2015.
- [7] Akhmad Arry.2008. *Proses Pembentukan dan Karakteristik Sinyal Ucapan*, Bandung : ITB .
- [8] D. Rhomanzah, "Sistem Kecerdasan Buatan Untuk Robot Asisten Berbasis Algoritma Case Base Reasoning," p. 8, 2015.
- [9] T. Chamidy, "Metode Mel Frequency Cepstral Coeffisients (MFCC) Pada klasifikasi Hidden Markov Model (HMM) Untuk Kata Arabic pada Penutur Indonesia," *Matics*, vol. 8, no. 1, pp. 36–39, 2016.
- [10] S. Dhingra, G. Nijhawan, and P. Pandit, "Isolated speech recognition using MFCC and DTW," *Int. J. Adv. ...*, vol. 2, no. 8, pp. 4085–4092, 2013.
- [11] A, Ixora. "Perancangan dan Implementasi Penerjemah Bahasa Isyarat dari Video Menjadi Suara Menggunakan Ekstraksi Ciri dan Hidden Markov Model". Tugas Akhir. Teknik Telekomunikasi. STT Telkom: Bandung.2010.
- [12] Rabiner, L.R., 1989, A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition, *Proceedings of the IEEE*, Vol. 77, No. 2, pp. 257-286.
- [13] T. Haryanto, "PENGUNAAN HIDDEN MARKOV MODEL ( HMM ) Untuk Mengidentifikasi RNA Family," Departemen Ilmu Komputer FMIPA IPB, 2010.